Testing Weak Instruments

If you have more than 1 Endog regressor, then the then the F-test for weak instruments is no longer valid.

In this instance, you use canonical correlations to assess instrument strength. Canonical correlations extend the correlation concept to <u>sets</u> of variables.

(A Bivariate, linear association)

Take 2 Endogenous regressors

$Y = \{Y_2, Y_3\}$

let $X_{EX} \equiv \{1, X_2, \ldots, X_G\}$

be the set of exogenous regressors.

let $\underset{\sim}{Z} = \{Z_1, Z_2\}$ be the external instruments.

let $M_x = I - X_{EX}(X_{EX}^T X_{EX})^{-1} X_{EX}^T$

Residual maker will remove the correlation with respect to Exog. Regressors.

$E_1 = M_X Y \quad n \times 2 \quad = (\hat{\underset{\sim}{e}}_{y2} \ \hat{\underset{\sim}{e}}_{y3})$

$E_2 = M_X Z \quad n \times 2 \quad = (\hat{\underset{\sim}{e}}_{z1} \ \hat{\underset{\sim}{e}}_{z2})$

$$Y_1^* = h_{11} \hat{e}_{y2} + h_{12} \hat{e}_{y3}$$

$$Z_1^* = k_{11} \hat{e}_{z1} + k_{12} \hat{e}_{z2}$$

$h_{11}, h_{12}, k_{11}, k_{12}$ are chosen to maximize correlation Between Y_1^* and Z_1^*

The resulting correlation is r_1 and is called the first canonical correlation.

Then

$$Y_2^* = h_{21}\hat{e}_{y2} + h_{22}\hat{e}_{y3}$$
$$Z_2^* = k_{21}\hat{e}_{z1} + k_{22}\hat{e}_{z2}$$

choose $h_{21}, h_{22}, k_{21}, k_{22}$ to maximize correlation between Y_2^* and Z_2^*, to produce r_2

If there are B variables in first group (Endog regressors) and L in second (external instruments)

$L \geq B$ for identification

There will be B possible canonical correlations

$$r_1 \geq r_2 \geq \cdots \geq r_B$$

The Cragg Donald F stat

$$F = \frac{\left[(N - G - B)/L\right]}{r_B^2 / (1 - r_B^2)}$$

critical values are in Stock & Yogo (2005)

N – sample size

G ~~B~~ – # exog regressors

B – # endog regressors

L – # external instruments

r_B^2 – smallest ~~co~~ canonical correlation.

LIML

Single equation estimator for use when the errors are normally distributed. Also used in other circumstances as it appears to be reasonably robust to other errors.

Let

$$Y^0 = \left[y_1 \; Y_2 \cdots Y_B \right]_{n \times (B+1)}$$

$$X_G = \left[1 \; x_2 \cdots x_G \right]_{n \times G}$$

$$X = \left[1 \; x_2 \cdots x_G \; z_1 \cdots z_L \right]_{n \times (G+L)}$$

Full set of exog vars including ext. instruments

Residual Maker $\quad M_X = I - X(X^T X)^{-1} X^T$

$$M_{X_G} = I - X_G (X_G^T X_G)^{-1} X_G^T$$

Resids $\quad E_0 = M_{X_G} \cdot Y^0 \qquad W_0 \equiv E_0^T E_0 \quad$ SSE

Resids $\quad E_1 = M_X Y^0 \qquad W_1 \equiv E_1^T E_1 \quad$ SSE

$$W \equiv (W_1)^{-1} W_0$$

and find the smallest characteristic root of W, λ_1

Partition W_0 and W_1

$$\begin{array}{ccc} & 1\times 1 & \\ W_0 = & \begin{bmatrix} w_{00} & \underset{\sim}{\omega}_0^T \\ \underset{\sim}{\omega}_0 & W_{00} \end{bmatrix} & \\ & B\times 1 \qquad\qquad B\times B & \end{array}$$

$$W_1 = \begin{bmatrix} w_{11} & \underset{\sim}{\omega}_1^T \\ \underset{\sim}{\omega}_1 & W_{11} \end{bmatrix} \; B\times B$$

Then

$$\hat{\gamma}_{LIML} = \left[W_{00} - \lambda_1 W_{11}\right]^{-1} \left(\omega_0^T - \lambda_1 \omega_1^T\right)$$

coeffs on the endog regressors.

$$\hat{\beta}_{LIML} = \left[X_0^T X_0\right]^{-1} X_0^T \left(y - \underset{-}{Y} \hat{\gamma}_{LIML}\right)$$

coeffs on the exog vars.

AS A k-class estimator.

$$\hat{\delta} = \begin{pmatrix} \hat{\gamma} \\ \hat{\beta} \end{pmatrix} = \begin{bmatrix} Y^T Y - k \hat{V}^T \hat{V} & Y^T X_G \\ X_G^T Y & X_G^T X_G \end{bmatrix} \circ$$

$$\begin{bmatrix} Y^T y - k \hat{V}^T \hat{v} \\ X_G^T y \end{bmatrix}$$

$\hat{V} = P_x Y$ residuals from red fm

$\hat{v} = P_x y$ residuals from red. F.

$k = 0$ — OLS

$k = 1$ — 2SLS

$k = \lambda_1$ — LIML

Anderson & Rubin Test

H_0: Eq is just identified

H_A: Eq is overidentified

$$LR = N(\lambda_1 - 1) \sim \chi^2_{L-B} \text{ is } H_0 \text{ True.}$$

Rejection implies that one or more of the omitted external instruments should not have been omitted

Note if $(W_1^{-1}) W_0 = I \Rightarrow \lambda_1 = 1$

$$\Rightarrow E_0^T E_0 = E_1^T E_1$$

External instruments Δ nothing.

```
open "@gretldir\data\poe\mroz.gdt"
square exper
series nwifeinc = (faminc-wage*hours)/1000
smpl hours>0 --restrict

matrix y1 = { hours, mtr, educ }
matrix w = { kidsl6, nwifeinc, const, exper, mothereduc, fathereduc}
matrix z = { kidsl6, nwifeinc, const}
matrix Mz = I($nobs)-z*invpd(z'*z)*z'
matrix Mw = I($nobs)-w*invpd(w'*w)*w'
matrix Ez= Mz*y1
matrix W0 = Ez'*Ez
matrix Ew = Mw*y1
matrix W1 = Ew'*Ew
matrix G = inv(W1)*W0
matrix l = eigengen(G, null)
scalar minl = min(l)
printf "\nThe minimum eigenvalue is %.8f \n",minl
matrix X = { mtr, educ, kidsl6, nwifeinc, const }
matrix y = { hours }
matrix kM = (I($nobs)-(minl*Mw))
matrix b =invpd(X'*kM*X)*X'*kM*y
a=rownames(b, " mtr educ kidsl6 nwifeinc const ")
printf "\nThe liml estimates are \n %.6f \n", b
```